## 5.3 Iterative Methods

A plethora of iterative methods has been proposed for the numerical solution of discrete models in tomography. However, only the algebraic reconstruction technique (ART) and EM have found widespread use. Thus we concentrate on these two methods and we survey the others in subsection 5.3.3.

### 5.3.1 ART

ART is simply the well-known Kaczmarz method for solving (over- or underdetermined) linear systems. Let $R_j f = g_j$ be such a system with $R_j : H \to H_j$, $j = 1, \ldots, p$, bounded linear operators from the Hilbert space $H$ onto (surjection) the Hilbert space $H_j$, $j = 1, \ldots, p$. We write

$$R = \begin{pmatrix} R_1 \\ \vdots \\ R_p \end{pmatrix}, \quad g = \begin{pmatrix} g_1 \\ \vdots \\ g_p \end{pmatrix}, \quad Rf = g.$$

The orthogonal projection $P_j$ in $H$ onto the affine subspace $R_j f = g_j$ is given by

$$P_j f = f + R_j^*(R_j R_j^*)^{-1}(g_j - R_j f). \tag{5.45}$$

We put for $\omega > 0$

$$P_j^\omega = (1-\omega)I + \omega P_j, \quad P^\omega = P_p^\omega \ldots P_1^\omega.$$

The Kaczmarz procedure with relaxation factor $\omega$ for solving $Rf = g$ is then

$$f^{k+1} = P^\omega f^k. \tag{5.46}$$

We describe one step explicitly. Putting $f^{k,0} = f^k$ and computing $f^{k,j}$, $j = 1, \ldots, p$, according to

$$f^{k,j} = P_j^\omega f^{k,j-1} = f^{k,j-1} + \omega R_j^*(R_j R_j^*)^{-1}(g_j - R_j f^{k,j-1}), \quad j = 1, \ldots, p, \tag{5.47}$$

we have $f^{k+1} = f^{k,p}$. For $\omega = 1$ and a linear system $Rf = g$ consisting of $p$ scalar equations in $n$ unknowns, this method was suggested by Kaczmarz (1937). With $a_1, \ldots, a_p$ the rows of $R$, it reads

$$f^{k,j} = f^{k,j-1} + \omega \frac{g_j - a_j f^{k,j-1}}{\|a_j\|^2} a_j^T, \quad j = 1, \ldots, p.$$

We work out the details of (5.47) for 2D transmission tomography in the standard parallel geometry (5.12). In that case,

$$(R_j f)(s) = (Rf)(\theta_j, s) = g_j(s)$$

## 5.3 Iterative Methods

A plethora of iterative methods has been proposed for the numerical solution of discrete models in tomography. However, only the algebraic reconstruction technique (ART) and EM have found widespread use. Thus we concentrate on these two methods and we survey the others in subsection 5.3.3.

### 5.3.1 ART

ART is simply the well-known Kaczmarz method for solving (over- or underdetermined) linear systems. Let $R_j f = g_j$ be such a system with $R_j : H \to H_j$, $j = 1, \ldots, p$, bounded linear operators from the Hilbert space $H$ onto (surjection) the Hilbert space $H_j$, $j = 1, \ldots, p$. We write

$$R = \begin{pmatrix} R_1 \\ \vdots \\ R_p \end{pmatrix}, \quad g = \begin{pmatrix} g_1 \\ \vdots \\ g_p \end{pmatrix}, \quad Rf = g.$$

The orthogonal projection $P_j$ in $H$ onto the affine subspace $R_j f = g_j$ is given by

$$P_j f = f + R_j^*(R_j R_j^*)^{-1}(g_j - R_j f). \tag{5.45}$$

We put for $\omega > 0$

$$P_j^\omega = (1-\omega)I + \omega P_j, \quad P^\omega = P_p^\omega \ldots P_1^\omega.$$

The Kaczmarz procedure with relaxation factor $\omega$ for solving $Rf = g$ is then

$$f^{k+1} = P^\omega f^k. \tag{5.46}$$

We describe one step explicitly. Putting $f^{k,0} = f^k$ and computing $f^{k,j}$, $j = 1, \ldots, p$, according to

$$f^{k,j} = P_j^\omega f^{k,j-1} = f^{k,j-1} + \omega R_j^*(R_j R_j^*)^{-1}(g_j - R_j f^{k,j-1}), \quad j = 1, \ldots, p, \tag{5.47}$$

we have $f^{k+1} = f^{k,p}$. For $\omega = 1$ and a linear system $Rf = g$ consisting of $p$ scalar equations in $n$ unknowns, this method was suggested by Kaczmarz (1937). With $a_1, \ldots, a_p$ the rows of $R$, it reads

$$f^{k,j} = f^{k,j-1} + \omega \frac{g_j - a_j f^{k,j-1}}{\|a_j\|^2} a_j^T, \quad j = 1, \ldots, p.$$

We work out the details of (5.47) for 2D transmission tomography in the standard parallel geometry (5.12). In that case,

$$(R_j f)(s) = (Rf)(\theta_j, s) = g_j(s)$$

with $R$ the 2D Radon transform, considered as an operator from $H = L_2(|x| < \rho)$ into $L_2(S^1 \times (-\rho, \rho))$. Thus $R_j$ is an operator from $H$ into $H_j = L_2(-\rho, \rho)$. From (2.8) we get

$$(R_j^* g)(x) = g(x \cdot \theta_j), \quad |x| < \rho,$$

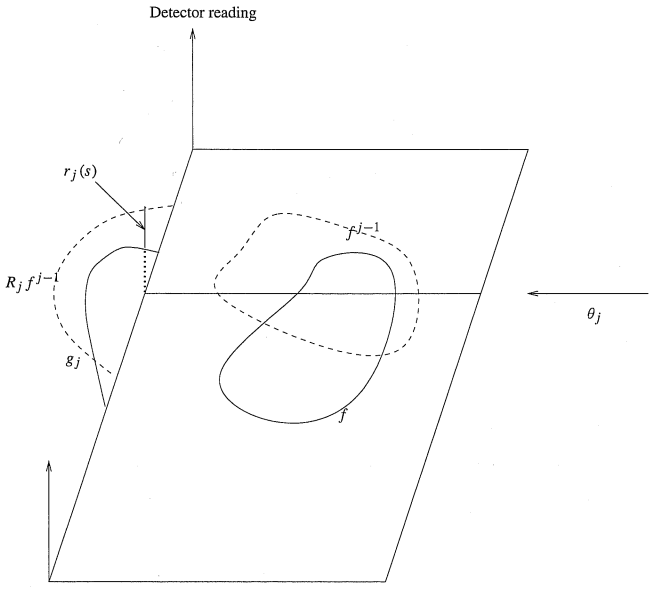


**Figure 5.6.** *ART for 2D transmission tomography.*

and hence

$$(R_j R_j^* g)(s) = 2\sqrt{\rho^2 - s^2}\, g(s), \quad |s| < \rho.$$

Thus (5.47) reads

$$\begin{aligned} f^{k,j}(x) &= f^{k,j-1}(x) + \frac{\omega}{2\sqrt{\rho^2 - (x \cdot \theta_j)^2}} r_j(x \cdot \theta_j), \\ r_j &= g_j - R_j f^{k,j-1}. \end{aligned} \tag{5.48}$$

This is easy to visualize; see Figure 5.6. $g_j$ is the projection of the true object $f$ as seen by the detector for direction $\theta_j$. $R_j f^{k,j-1}$ is the projection of the current approximation $f^{k,j-1}$. The discrepancy $r_j(s) = g_j(s) - (R_j f^{k,j-1})(s)$ at detector position $s$ must come from a mismatch of $f$ and $f^{k,j-1}$ along the line $x \cdot \theta_j = s$. There is no information whatsoever concerning where on this line the mismatch occurs. Thus all we can do to correct the mismatch is to spread the discrepancy $r_j(s)$ evenly about the line $x \cdot \theta_j = s$, possibly with a suitable weight. This is exactly what (5.48) is doing.

We carry out the convergence analysis for a method slightly more general than (5.47). We replace the operator $R_j R_j^*$ of that formula by an arbitrary positive definite operator $C_j$ in $H_j$. This yields

$$\begin{aligned} f^{k,j} &= f^{k,j-1} + \omega R_j^* C_j^{-1}(g_j - R_j f^{k,j-1}), \quad j = 1, \ldots, p, \\ f^{k,0} &= f^k, \quad f^{k+1} = f^{k,p}. \end{aligned} \tag{5.49}$$

Special cases are the Landweber iteration ($p = 1$, $C_1 = I$; see Hanke, Neubauer, and Scherzer (1995)) and fixed-block ART (dim $H_j$ finite, $C_j$ diagonal; see Censor and Zenios (1997)). We consider only the consistent case in which $Rf = g$ has a solution.

**Theorem 5.1.** *Let $C_j$ be bounded and positive definite, and let $C_j \geq R_j R_j^* > 0$, $j = 1, \ldots, p$. Let $Rf = g$ be consistent. Then the iteration* (5.49) *is convergent, and*

$$f^k \to P_R f^0 + R^+ g,$$

*where $P_R$ is the orthogonal projection on* $\ker(R)$ *and $R^+$ is the generalized inverse of $R$ (see section* 1.3.3).

***Proof.*** For $C_j = R_j R_j^*$ this is essentially Theorem 3.9 of Natterer (1986). We give a more elementary proof along the lines of Tanabe (1971).

Let $f^+$ be the minimal norm solution of $Rf = g$, i.e., $Rf^+ = g$ and $f^+ \in \ker(R)^\perp$, and let $e^k = f^k - f^+$. Then

$$\begin{aligned} e^{k+1} &= Qe^k, \quad Q = Q_p \cdots Q_1, \\ Q_j &= I - \omega R_j^* C_j^{-1} R_j, \quad j = 1, \ldots, p. \end{aligned}$$

$Q$ has the invariant subspaces $\ker(R)$ and $\ker(R)^\perp$, and $Q = I$ on $\ker(R)$. We show that $Q^k \to 0$ as $k \to \infty$ strongly on $\ker(R)^\perp$, proving the theorem.

We have

$$\begin{aligned} \|Q_j f\|^2 &= \|f\|^2 - 2\omega(f, R_j^* C_j^{-1} R_j f) + \omega^2 (R_j^* C_j^{-1} R_j f, R_j^* C_j^{-1} R_j f) \\ &= \|f\|^2 - 2\omega(R_j f, C_j^{-1} R_j f) + \omega^2 (R_j R_j^* C_j^{-1} R_j f, C_j^{-1} R_j f). \end{aligned}$$

Since $R_j R_j^* \leq C_j$, we have

$$\begin{aligned} (R_j R_j^* C_j^{-1} R_j f, C_j^{-1} R_j f) &\leq (C_j C_j^{-1} R_j f, C_j^{-1} R_j f) \\ &= (R_j f, C_j^{-1} R_j f), \end{aligned}$$

and hence

$$\begin{aligned} \|Q_j f\|^2 &\leq \|f\|^2 - 2\omega(R_j f, C_j^{-1} R_j f) + \omega^2 (R_j f, C_j^{-1} R_j f) \\ &= \|f\|^2 - \omega(2 - \omega)(R_j f, C_j^{-1} R_j f). \end{aligned} \tag{5.50}$$

Since $C_j$ is positive definite and $0 < \omega < 2$, we see that $\|Q_j f\| \leq \|f\|$, with equality only for $R_j f = 0$. It follows that $\|Qf\| \leq \|f\|$, with equality only for $f \in \ker(R)$. If $0 \neq f \in \ker(R)^\perp$, then $\|Qf\| < \|f\|$, because otherwise $f \in \ker(R)$, and hence $f = 0$. For $\ker(R)^\perp$ finite dimensional, it follows that $Q$ is a contraction on $\ker(R)^\perp$, and hence $Q^k \to 0$ strongly on $\ker(R)^\perp$. In the infinite dimensional case, this follows from the following fact: Let $Q$ be a self-adjoint operator in a Hilbert space with $\|Q\| \leq 1$ and $\|Qf\| < \|f\|$ for $f \neq 0$. Then $Q^k \to 0$ strongly as $k \to \infty$.

For the proof of this fact, we first observe that the sequence $\varepsilon_k = \|Q^k f\|$ is nonincreasing and hence convergent. It suffices to show that its limit $\varepsilon$ is zero. We have

$$\begin{aligned}\|Q^{2k} f - Q^{2k+2m} f\|^2 &= \|Q^{2k} f\|^2 + \|Q^{2k+2m} f\|^2 - 2(Q^{2k+m} f, Q^{2k+m} f) \\ &= \varepsilon_{2k}^2 + \varepsilon_{2k+2m}^2 - 2\varepsilon_{2k+m}^2 \to 0\end{aligned}$$

as $k \to \infty$ for each $m$, and hence $Q^{2k} f \to g$ for some $g$ as $k \to \infty$. It follows that

$$\begin{aligned}\|g\| &= \lim_k \|Q^{2k} f\| = \varepsilon, \\ \|Qg\| &= \lim_k \|Q^{2k+1} f\| = \varepsilon,\end{aligned}$$

and hence $\|Qg\| = \|g\|$. This is possible only for $g = 0$. Hence $\varepsilon = 0$. This establishes $Q^k \to 0$ strongly in $\ker(R)^\perp$.

To finish the proof, we write

$$e^0 = f^0 - f^+ = P_R f^0 + (I - P_R) f^0 - f^+.$$

Since $Q = I$ on $\ker(R)$ and $f^+$, $(I + P_R) f^0 \in \text{range}(R^*)$, we have

$$f^k - f^+ = e^k = Q^k e^0 = P_R f^0 + Q^k((I - P_R) f^0 - f^+) \to P_R f^0$$

as $k \to \infty$. □

As a by-product the proof provides interesting information about the possible gain in accuracy in one step of the iteration. From (5.50) we get

$$\begin{aligned}\|f^{k,j} - f\|^2 \le \|f^{k,j-1} - f\|^2 \\ - \omega(2 - \omega)(R_j f^{k,j-1} - g_j, C_j^{-1}(R f^{k,j-1} - g_j)),\end{aligned} \tag{5.51}$$

where $f = P_R f^0 + R^+ g$ is the limit. Thus the error becomes smaller whenever the residual $R_j f^{k,j-1} - g_j$ does not vanish, and the amount of the improvement can easily be computed.

Theorem 5.1 is clearly reminiscent of the SOR theory for linear systems; see Ortega and Rheinboldt (1970). In fact, there are close connections between the Kaczmarz method for $Rf = g$ and the SOR method for $RR^*u = g$. More precisely, we have the next theorem.

**Theorem 5.2.** *Let $u^k$ be the SOR iterates for $RR^*u = g$, i.e.,*

$$\begin{aligned}u^{k+1} &= C_\omega u^k + c_\omega, \\ C_\omega &= I - \omega(D + \omega L)^{-1} RR^*, \quad c_\omega = \omega(D + \omega L)^{-1} g.\end{aligned}$$

*Then $f^k = R^* u^k$ are the Kaczmarz iterates* (5.46) *for $Rf = g$.*

So far we have considered the pure linear algebra point of view. In order to learn more about the speed of convergence and the qualitative behavior of the iterates, we exploit the analytic structure of the problems in tomography. We consider the case of reconstruction from complete projections. Let $H = L_2(B)$, let $B$ be the unit ball in $\mathbb{R}^n$, and let $H_j = L_2(-1, +1; w^{1-n})$ with $w(s) = (1 - s^2)^{1/2}$. The maps $R_j : H \to H_j$ are defined by

$$(R_j f)(s) = (Rf)(\theta_j, s), \quad j = 1, \ldots, p,$$

with $R$ the Radon transform (2.1), are bounded and surjective. The adjoint $R_j^* : H_j \to H$ is given by

$$(R_j^* g)(x) = (w^{1-n} g)(x \cdot \theta_j);$$

compare (2.7). A straightforward computation shows that

$$R_j R_j^* = \frac{1}{n-1} |S^{n-2}| I.$$

Hence (5.45) reads

$$P_j f(x) = f(x) + \frac{n-1}{|S^{n-2}|} (w^{1-n}(g_j - R_j f))(x \cdot \theta_j), \tag{5.52}$$

and the ART iteration (5.47) assumes the form

$$f^{k,j}(x) = f^{k,j-1}(x) + \omega'(w(g_j - R_j f^{k,j-1}))(x \cdot \theta_j), \quad \omega' = \frac{(n-1)\omega}{|S^{n-2}|}.$$

From Theorem 5.1 we know that, in the consistent case, $f^k \to f^+$, where $f^+$ is the solution of minimal norm. The errors $e^k = f^+ - f^k$ satisfy

$$e^k = Q^\omega e^{k-1}, \quad Q^\omega = Q_p^\omega \cdots Q_1^\omega, \quad Q_j^\omega = (1-\omega) I + \omega Q_j,$$

where $Q_j$ is the orthogonal projection onto the subspace $R_j f = 0$. We obtain $Q_j$ from (5.52) with $g_j = 0$, i.e.,

$$Q_j f(x) = f(x) - \frac{n-1}{|S^{n-2}|} (w^{1-n} R_j f)(x \cdot \theta_j).$$

Let $C_m^\lambda$, $\lambda > -1/2$, be the Gegenbauer polynomial of degree $m$, i.e., the orthogonal polynomials on $[-1, +1]$ with weight $(1-s^2)^{\lambda - 1/2}$; see section 1.3.5. We put

$$C_{m,j}(x) = C_m^{n/2}(x \cdot \theta_j), \quad \mathcal{C}_m = \langle C_{m,1}, \ldots, C_{m,p} \rangle.$$

**THEOREM 5.3.** *Let* $\alpha_m(t) = \frac{\pi^{(n-1)/2}}{\Gamma(\frac{n+1}{2})} \frac{C_m^{n/2}(t)}{C_m^{n/2}(1)}$. *Then*

$$Q_j C_{m,i} = C_{m,i} - \frac{n-1}{|S^{n-2}|} \alpha_m(\theta_i \cdot \theta_j) C_{m,j}.$$

The theorem was obtained by Hamaker and Solmon (1978). It implies that $\mathcal{C}_m$ is an invariant subspace of $Q_j$ and hence of $Q_j^\omega$ and of $Q^\omega$. We compute a matrix representation of $Q^\omega$ in $\mathcal{C}_m$. One can show that

$$\dim \mathcal{C}_m = \min\{p, M\}, \quad M = \begin{pmatrix} m+n-1 \\ n-1 \end{pmatrix}, \tag{5.53}$$

provided that

$$\theta_j \neq \pm \theta_k, \quad k \neq j, \quad k, j = 1, \ldots, p. \tag{5.54}$$

For $M > p$, convergence on $\mathcal{C}_m$ is of no interest since for these $m$ the functions in $\mathcal{C}_m$ represent details that, by our resolution analysis in section 4.2, cannot be resolved anyway. Therefore, we assume in the following that $M \leq p$.

As a direct consequence of Theorem 5.3, $Q_j^\omega$ in $\mathcal{C}_m$ is represented by the $p \times p$ matrix

$$A_{m,j}(\omega) = \begin{pmatrix} 1 & & & & & & \\ & 1 & & & & & \\ -\omega\alpha_{m,j,1} & \cdots & 1-\omega & \cdots & -\omega\alpha_{m,j,p} & & \\ & & & 1 & & & \\ & & & & & 1 & \end{pmatrix}, \quad \alpha_{m,k,j} = \frac{n-1}{|S^{n-2}|}\alpha_m(\theta_k \cdot \theta_j),$$

which differs from the unit matrix only in row $j$. For $n = 2$ and $\theta_j = (\cos\varphi_j, \sin\varphi_j)^T$ we have $M = m + 1$ and

$$\alpha_{m,k,j} = \begin{cases} \frac{\sin((m+1)(\varphi_k - \varphi_j))}{(m+1)\sin(\varphi_k - \varphi_j)}, & \varphi_k \neq \varphi_j, \\ 1, & \varphi_k = \varphi_j. \end{cases}$$

The matrix representation of $Q^\omega$ in $\mathcal{C}_m$ is

$$A_m(\omega) = A_{m,p}(\omega) \cdots A_{m,1}(\omega). \tag{5.55}$$

One can also show that

$$\text{range}(R^*) = \bigoplus_{m=0}^{\infty} \mathcal{C}_m.$$

For $f^0 \in \text{range}(R^*)$, all $e^k$ are in $\text{range}(R^*)$. Thus we can study the speed of convergence for each $\mathcal{C}_m$ separately simply by computing the spectral radius $\rho_m(\omega)$ of $Q^\omega$ on $\mathcal{C}_m$ for all $m$ with $M \leq p$. This is facilitated by the fact that the inner products of the $C_{m,i}$ are known: With some constant $c(n, m)$ we have

$$\int_{|x|<1} C_{m,k}(x) C_{m,j}(x) dx = c(n, m) C_m^{n/2}(\theta_k \cdot \theta_j).$$

Numerical results for $\rho_m(\omega)$ are presented in Figure 5.7 for $n = 2$ and $p = 30$ directions $\theta_j$.

The crucial point is that the $\rho_m(\omega)$ depend decisively on the way in which the directions $\theta_j$ are ordered. First, we consider the sequential order, i.e., $\varphi_j = \frac{j-1}{p}\pi$, $j = 1, \ldots, p$. In Figure 5.7(a) we see that for $\omega = 1$, convergence is fast on the $\mathcal{C}_m$ with $m$ large, i.e., on the high-frequency components, while convergence is slow on the $\mathcal{C}_m$ with $m$ small, i.e., on the low-frequency components. The situation is just the other way around for $\omega = 0.1$. Thus for $\omega = 1$, the high-frequency components (such as noise) show up early in the iteration, while overall features are determined later. It is clear that this is not a desirable behavior. For $\omega = 0.1$ the iterations first determine the smooth parts of $f$ and the small details only later. This is clearly more desirable. Thus surprisingly small values of $\omega$ (e.g., $\omega = 0.05$) are quite common in ART.

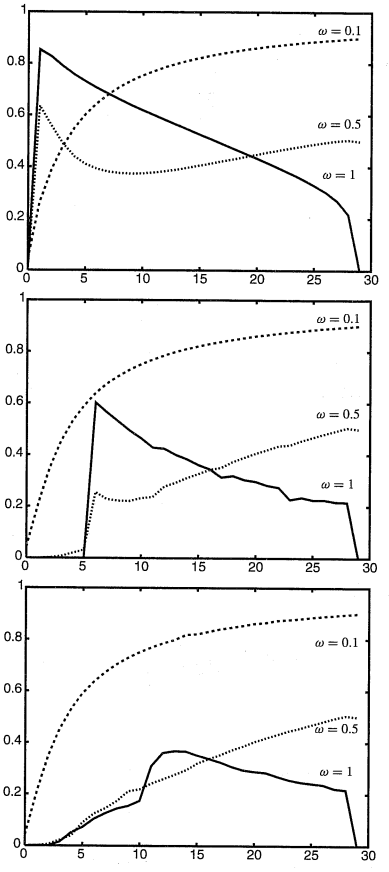


**Figure 5.7.** *Spectral radius $\rho_m(\omega)$ of $Q^\omega$ in $\mathcal{C}_m$ for $n = 2$, $p = 30$ as a function of $m$.* (a) *Sequential order of directions.* (b) *Nonsequential order.* (c) *Random order.*

The situation changes drastically if the linear order of directions is given up in favor of a nonsequential order; see Figure 5.7(b). For $\omega = 1$ we now have fast convergence on all subspaces $\mathcal{C}_m$, i.e., for all spatial frequencies. The practical consequence is that it is advisable to use an order different from the sequential one. This was discovered by Hamaker and Solmon (1978) and was rediscovered in Herman and Meyer (1993). A good strategy for choosing an order is to make directions as orthogonal as possible to the previous ones. For example, for 18 parallel projections in the plane, the directions 0, 90, 140, 50, 110, 30, 160, 70, 130, 20, 100, 170, 40, 80, 150, 10, 60, 120 degrees seem to be favorable. In fact, Hamaker and Solmon (1978) showed that they are the best possible for $p = 18$. However, a random choice of directions is almost as good; see Figure 5.7(c).

It is interesting to compare ART with other iterative methods, such as the Landweber iteration, which came to be known as SIRT (simultaneous iterative reconstruction technique) in tomography. The difference from (5.52) is that the update is done only after a complete sweep through all the directions. This leads to

$$f^{k+1} = f^k + \omega \sum_{j=1}^{p} R_j^* C_j^{-1} (g_j - R_j f^k).$$

Introducing $e^k = f^k - f^+$, we get

$$e^{k+1} = \overline{Q}^{\omega} e^k,$$

where now, for $C_j = R_j R_j^*$,

$$\begin{aligned} \overline{Q}^{\omega} &= I - \omega \sum_{j=1}^{p} R_j^* C_j^{-1} R_j \\ &= I + \omega \sum_{j=1}^{p} (Q_j - I) \end{aligned}$$

with $Q_j$ as above. Thus by Theorem 5.3, $\mathcal{C}_m$ is again an invariant subspace of $\overline{Q}^{\omega}$, and the matrix representation of $\overline{Q}^{\omega}$ is

$$\overline{A}_m(\omega) = I - \omega \alpha_m,$$

where $\alpha_m$ is the $(p, p)$-matrix with elements $\alpha_{m,i,j}$, $i, j = 1, \ldots, p$. For evenly spaced directions $\theta_j$, i.e., $\varphi_j = (j-1)\pi/p$, $j = 1, \ldots, p$, the eigenvalues of $\alpha_m$ are easy to compute. Using the identity

$$(m+1)\alpha_{m,k,j} = e^{im\psi} + e^{i(m-2)\psi} + \cdots + e^{-im\psi},$$

where $\psi = \pi(k-j)/p$, the eigenvalue equation $\alpha_m x = \lambda x$, $x = (x_0, \ldots, x_{p-1})^T$, reads

$$(m+1) \sum_{j=0}^{p-1} {\sum_{\ell=-m}^{m}}' e^{i\pi\ell(k-j)/p} x_j = \lambda x_k, \quad k = 0, \ldots, p-1,$$

where the prime at the $\ell$-sum indicates that $\ell + m$ is even. Putting $\ell + m = 2\nu$, rearranging yields

$$(m+1) \sum_{\nu=0}^{m} e^{2\pi i \nu k/p} \sum_{j=0}^{p-1} e^{-2\pi i \nu j/p} y_j = \lambda y_k, \quad k = 0, \ldots, p-1,$$

where $y_j = e^{i\pi m j/p} x_j$, $j = 0, \ldots, p-1$. With the matrix $F = (e^{-2\pi i jk/p})_{j,k=0,\ldots,p-1}$ and $y = (y_0, \ldots, y_{p-1})^T$, this reads

$$\frac{m+1}{p} F^{-1} P F y = \lambda y,$$

where $P$ is the diagonal matrix with $m+1$ 1's and $p-m-1$ 0's on the diagonal. It follows that $\alpha_m$ has the eigenvalue $\frac{p}{m+1}$ with multiplicity $m+1$ and 0 with multiplicity $p-m-1$. Thus the spectral radius $\overline{\rho_m}(\omega)$ of $\overline{Q}^{\omega}$ on $\mathcal{C}_m$ is

$$\overline{\rho_m}(\omega) = \left| 1 - \frac{\omega p}{m+1} \right|.$$

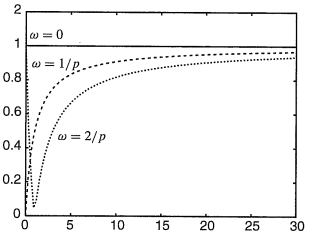


**Figure 5.8.** *Spectral radius* $\overline{\rho_m}(\omega)$ *of* $\overline{Q}^{\omega}$ *in* $C_m$ *for* $n = 2,\ p = 30$.

Hence the condition for convergence of SIRT is $0 < \omega < \frac{2}{p}$.

Comparing the graph of $\overline{\rho_m}(\omega)$ in Figure 5.8 with $\rho_m(\omega)$ in Figure 5.7 makes clear why ART (with a good arrangement of directions or a favorable choice of $\omega$) is so much better than SIRT.

In practice, ART is applied to a discretized version of the linear integral equation governing a certain imaging system, e.g., the Radon integral equation. One represents the unknown function $f$ by an expansion of the form

$$f = \sum_{k=1}^{N} f_k B_k$$

with certain basis functions $B_k$. These basis functions can be chosen, for example, as the characteristic functions of pixels or voxels. It has been suggested to use radially symmetric functions, the so-called "blobs"; see Lewitt (1992) and Marabini et al. (1999). With $g_i$ the $i$th measurements, we then obtain

$$g_i = \sum_{k=1}^{N} a_{ik} f_k,$$

where $a_{ik}$ is the $i$th measurement of $B_k$. In imaging, this approach is sometimes called a series expansion method, while in numerical analysis, one would call it a projection method, or more specifically a collocation method.

### 5.3.2 The EM algorithm

This is an iterative algorithm for maximizing the likelihood function $L$ from (3.9). Taking the logarithm of $L$ and omitting an additive constant, we can as well maximize the log likelihood function

$$\ell(f) = \sum_{i=1}^{n} (g_i \log(Af)_i - (Af)_i). \tag{5.56}$$

In order to avoid purely technical difficulties, we assume $A,\ g > 0$ elementwise. One can easily verify that the Hessian of $\ell$ is negative semidefinite, i.e., $\ell$ is concave. Thus local maxima of $\ell$ in $f \geq 0$ are also global ones, and $f$ is a global maximum if and only if the Kuhn–Tucker conditions